



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/058,684	01/28/2002	Baoxin Li	KLR 7146.125	4058
47915	7590	04/19/2005	EXAMINER	
CHERNOFF, VILHAUER, MCCLUNG & STENZEL, LLP 1600 ODS TOWER 601 SW SECOND AVENUE PORTLAND, OR 97204			ZHOU, TING	
		ART UNIT		PAPER NUMBER
				2173

DATE MAILED: 04/19/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)
	10/058,684	LI, BAOXIN
	Examiner	Art Unit
	Ting Zhou	2173

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) Responsive to communication(s) filed on 24 January 2005.
- 2a) This action is **FINAL**. 2b) This action is non-final.
- 3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) Claim(s) 1-64 is/are pending in the application.
- 4a) Of the above claim(s) 50-60 and 62-64 is/are withdrawn from consideration.
- 5) Claim(s) _____ is/are allowed.
- 6) Claim(s) 1-49 and 61 is/are rejected.
- 7) Claim(s) _____ is/are objected to.
- 8) Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) The specification is objected to by the Examiner.
- 10) The drawing(s) filed on 28 January 2002 is/are: a) accepted or b) objected to by the Examiner. Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a). Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
 - a) All b) Some * c) None of:
 1. Certified copies of the priority documents have been received.
 2. Certified copies of the priority documents have been received in Application No. _____.
 3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- 1) Notice of References Cited (PTO-892)
- 2) Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
Paper No(s)/Mail Date 7/26/02
- 4) Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____
- 5) Notice of Informal Patent Application (PTO-152)
- 6) Other: _____

DETAILED ACTION

1. Claims 1-64 are pending.
2. Applicants' response filed on 24 January 2005 have been received and entered.

Election/Restrictions

3. The applicants believe that the restriction should be partially traversed by combining Inventions I and II into a single application, while reserving Inventions III and IV for subsequent divisional applications. The examiner has fully considered the applicants' arguments and agrees with the applicants' arguments. The examiner combines previous inventions I & II, set forth in the restriction requirement mailed on 21 December 2004. The current grouping of the claims stands as:

Group I: Claims 1-49 and 61
Group II: Claims 50-54 and 62-64
Group III: Claims 55-60

4. The applicants have elected Group I (Claims 1-49 and 61) for prosecution; Claims 50-60 and 62-64 are withdrawn from consideration.
5. Prosecution on the merits of elected claims 1-49 and 61 follows:

6. Applicant is advised that should claim 20 be found allowable, claim 21 will be objected to under 37 CFR 1.75 as being a substantial duplicate thereof. When two claims in an application are duplicates or else are so close in content that they both cover the same thing, despite a slight difference in wording, it is proper after allowing one claim to object to the other as being a substantial duplicate of the allowed claim. See MPEP § 706.03(k).

Double Patenting

The nonstatutory double patenting rejection is based on a judicially created doctrine grounded in public policy (a policy reflected in the statute) so as to prevent the unjustified or improper timewise extension of the "right to exclude" granted by a patent and to prevent possible harassment by multiple assignees. See *In re Goodman*, 11 F.3d 1046, 29 USPQ2d 2010 (Fed. Cir. 1993); *In re Longi*, 759 F.2d 887, 225 USPQ 645 (Fed. Cir. 1985); *In re Van Ornum*, 686 F.2d 937, 214 USPQ 761 (CCPA 1982); *In re Vogel*, 422 F.2d 438, 164 USPQ 619 (CCPA 1970); and, *In re Thorington*, 418 F.2d 528, 163 USPQ 644 (CCPA 1969).

A timely filed terminal disclaimer in compliance with 37 CFR 1.321(c) may be used to overcome an actual or provisional rejection based on a nonstatutory double patenting ground provided the conflicting application or patent is shown to be commonly owned with this application. See 37 CFR 1.130(b).

Effective January 1, 1994, a registered attorney or agent of record may sign a terminal disclaimer. A terminal disclaimer signed by the assignee must fully comply with 37 CFR 3.73(b).

7. Claims 1-17 and 35-49 are provisionally rejected under the judicially created doctrine of obviousness-type double patenting as being unpatentable over claims 1-12, 27, 29, 36-41 and 46-50 of copending Application No. 09/934,004. Although the conflicting claims are not identical, they are not patentably distinct from each other because both sets of claims deal with methods of processing a video directed to a sports match including identifying a plurality of segments of the video based upon an event, wherein the event is characterized by a start time and an end time, and creating a summarization of the sports video by including the plurality of segments.

Although the start and end times in claims 1-17 and 35-49 of the present application is based upon events occurring for use in a video of a sumo match and the start and end times in claims 1-12, 27, 29, 36-41 and 46-50 of copending Application No. 09/934,004 is based upon the intended use of events for a video of a baseball game, both applications deal with creating a video summary from identified segments of a video, based upon sports event with a start and end time. The segments identified in the claimed limitations can be performed by the user and applicable to any video subject matter.

This is a provisional obviousness-type double patenting rejection because the conflicting claims have not in fact been patented.

Specification

8. Applicant is reminded of the proper content of an abstract of the disclosure.

A patent abstract is a concise statement of the technical disclosure of the patent and should include that which is new in the art to which the invention pertains. If the patent is of a basic nature, the entire technical disclosure may be new in the art, and the abstract should be directed to the entire disclosure. If the patent is in the nature of an improvement in an old apparatus, process, product, or composition, the abstract should include the technical disclosure of the improvement. In certain patents, particularly those for compounds and compositions, wherein the process for making and/or the use thereof are not obvious, the abstract should set forth a process for making and/or use thereof. If the new technical disclosure involves modifications or alternatives, the abstract should mention by way of example the preferred modification or alternative.

The abstract should not refer to purported merits or speculative applications of the invention and should not compare the invention with the prior art.

The abstract should be in narrative form and generally limited to a single paragraph on a separate sheet within the range of 50 to 150 words. It is important that the abstract not exceed 150 words in length since the space provided for the abstract on the computer tape used by the printer is limited. The form and legal phraseology often used in patent claims, such as "means" and "said," should be avoided. The abstract should describe the disclosure sufficiently to assist readers in deciding whether there is a need for consulting the full patent text for details.

The language should be clear and concise and should not repeat information given in the title. It should avoid using phrases which can be implied, such as, "The disclosure concerns," "The disclosure defined by this invention," "The disclosure describes," etc.

9. The abstract of the disclosure is objected to because it is not descriptive enough to sufficiently assist readers in deciding whether there is a need for consulting the full patent text for details; furthermore, the abstract does not include any description regarding that which is new in the art to which the invention pertains. Correction is required. See MPEP § 608.01(b).

Claim Rejections - 35 USC § 101

35 U.S.C. 101 reads as follows:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

10. Claims 1, 7, 11, 18, 22, 27, 31, 34, 35, 36, 39, 42, 45, 46, 48, 50, 61 and 62 are rejected under 35 U.S.C. 101 because the claimed invention is directed to a non-statutory abstract idea. The cited claims are directed to a video processing method, however, they fail to produce a practical application in the technological arts. For such subject matter to be statutory, the claimed process must be limited to a practical application of the abstract idea or mathematical algorithm in the technological arts. See Alappat, 33 F.3d at 1543, 31 USPQ2d at 1556-57 (quoting Diamond v. Diehr, 450 U.S. at 192, 209 USPQ at 10). See MPEP 2106. The claimed invention is directed solely to an abstract idea or to manipulation of abstract ideas, and does not produce a practical application in the technological arts. The steps recited in the claims, i.e. identify a plurality of segments of a video, and create a summarization of the video from the identified segments, can be done by a user in his mind, i.e. identify a plurality of segments of a video that the user watched by remembering a plurality of memorable segments of the video and combine the segments he remembers to create a mental summarization of the video; none of the claimed steps are required to be performed on or by a computer, or computer-related system, and amount to an abstract idea that does not produce a practical application in the technological arts.

11. To expedite a complete examination of the instant application the claims rejected under 35 U.S.C. 101 (nonstatutory) above are further rejected as set forth below in anticipation of applicant amending these claims to place them within the four statutory categories of invention.

Claim Rejections - 35 USC § 112

The following is a quotation of the second paragraph of 35 U.S.C. 112:

The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

12. Claims 42-44 and 61 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

Claim 42 recites the limitation “said summarization” in line 7 of the claim. There is insufficient antecedent basis for this limitation in the claims.

Claims 43-44 recite the limitation "said connecting" in line 1 of the claims. There is insufficient antecedent basis for this limitation in the claim. The examiner assumes that this is a typographical error and that for examination purposes, claims 43-44 were meant to be dependent upon claim 39, which has antecedent basis for “said connecting”, instead of claim 42.

Claim 61 recites the limitations “the detection” and “the center” on lines 2 and 5, respectively, of the claim. There is insufficient antecedent basis for these limitations in the claims.

Claim Rejections - 35 USC § 102

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

13. Claims 22, 27-28 and 31-33 are rejected under 35 U.S.C. 102(b) as being anticipated by “Event Detection and Summarization in Sports Video”, Li et al. (hereinafter Li).

Referring to claim 22, Li teaches a method comprising identifying a plurality of segments of a video, wherein the start of the plurality of segments is identified based upon a pair of regions having a dominant color description representative of skin tone, where each of the segments includes a plurality of frames of the video (Case Study 3: Japanese Sumo Wrestling; detecting a start scene based upon frames containing blobs of skin color); and creating a summarization of the video by including the plurality of segments, where the summarization includes fewer frames than the video (Abstract; concatenating detected plays to generate a compact summarization of the video).

Referring to claim 27, Li teaches a method comprising identifying a plurality of segments of a video, wherein the start of the plurality of segments is identified based upon a pair of regions generally symmetric to each other with respect to a generally center column of a frame of the video, where each of the segments includes a plurality of frames of the video (Case Study 3: Japanese Sumo Wrestling; detecting a start scene based upon frames containing two symmetrically distributed regions, or blobs of skin color); and creating a summarization of the video by including the plurality of segments, where the summarization includes fewer frames than the video (Abstract; concatenating detected plays to generate a compact summarization of the video).

Referring to claim 28, Li teaches the pair of spatial regions having a dominant color description representative of skin tone video (Case Study 3: Japanese Sumo Wrestling; detecting

a start scene based upon frames containing two symmetrically distributed regions, or blobs of skin color).

Referring to claim 31, Li teaches a method comprising identifying a plurality of segments of a video, wherein the start of the plurality of segments is identified based upon a pair of spatial regions that move toward one another, where each of the segments includes a plurality of frames of the video (Case Study 3: Japanese Sumo Wrestling; detecting a start scene based upon frames containing two symmetrically distributed blobs of skin color that converge); and creating a summarization of the video by including the plurality of segments, where the summarization includes fewer frames than the video (Abstract; concatenating detected plays to generate a compact summarization of the video).

Referring to claim 32, Li teaches the pair of spatial regions has a dominant color description representative of skin color (Case Study 3: Japanese Sumo Wrestling; detecting a start scene based upon frames containing two symmetrically distributed blobs of skin color that converge).

Referring to claim 33, Li teaches the pair of spatial regions collides with one another (Case Study 3: Japanese Sumo Wrestling; tracking the two blobs of skin color to see if they converge, i.e. collide with one another).

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person

having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

14. Claims 1-7, 9-17 and 36-37 are rejected under 35 U.S.C. 103(a) as being unpatentable over "Indexing of Baseball Telecast for Content-based Video Retrieval", Kawashima et al. (hereinafter Kawashima).

Referring to claim 1, Kawashima teaches a method of processing a video including baseball comprising: (a) identifying a plurality of segments of the video based upon an event, wherein the event is characterized by a start time based upon when the ball is put into play and an end time based upon when the ball is considered out of play, where each of the segments includes a plurality of frames of the video (pp. 871-873, sections 1.1, 1.2, 2.1 and 2.2; e.g. the at bat event comprising of a start point in time slightly before the pitching and endpoint in time slightly after the catcher catches the ball if the ball is struck out and after the ball is thrown to a baseman if the ball is hit); and

(b) creating a summarization of the video by including the plurality of segments, where the summarization includes fewer frames than the video (Abstract; pg. 872, section 1.2; i.e. the indexed video segments is a digest of the game or summary of the video, a.k.a. compressed play).

Although Kawashima does not explicitly teach the start time based upon specific sumo events such as when the players line up to charge one another and an end time based upon specific sumo events such as when one of the players at least one of steps outside the ring and touches the ring surface with part of his body other than the shoes of his feet, Kawashima teaches the start time and end time based upon baseball events. Since both sumo wrestling and baseball belong to a class of sporting events modeled as a sequence of "plays" identified by a start time

and end time, it would have been obvious to one of ordinary skill in the art to apply the video summarization of identified segments based upon baseball events taught by Kawashima to segments based upon other sporting events such as a sumo wrestling match. One would have been motivated to make such a combination in order to generate a compact representation of the content of video material, allowing easy browsing, filtering, indexing and retrieval, etc., enabling users only interested in the exciting highlights of a game to skip the long and often boring portions of watching a sporting match in its entirety.

In a similar manner, all subsequent claims dealing with teachings of limitations for use in a baseball game can be applied for use in any other sports games, including sumo wrestling.

Referring to claim 2, Kawashima teaches a method of processing a video including baseball wherein the event is defined by the rules of baseball (pp. 871-873, sections 1.1-2.1.4; events such as scenes in which a batter was struck out or got a hit or a home run is defined by the rules of baseball using a spotting technique comprising a search of the minimal warp function by comparing input video sequence with pitching/batting model sequences). Similarly, it would have been obvious that the event can be defined by the rules of sumo.

Referring to claims 3, 6 and 15, Kawashima teaches a method of processing a video including baseball wherein the start time is temporally proximate a baseball pitch (pg. 872, lines 10-11). Similarly, it would have been obvious that the start time can be temporally proximate a sumo event, such as a charge of the two players, or includes a portion of the pre-bout ceremonies.

Referring to claims 4, 5, 16 and 17, Kawashima teaches a method of processing a video including baseball wherein the end time is temporally proximate to the batter missing the ball

with a bat (pg. 872, lines 12-15). Similarly, it would have been obvious that the end time can be temporally proximate a sumo event, such as stepping out of the ring or touching of the surface of the ring by a part of his body other than the soles of his feet.

Referring to claim 7, Kawashima teaches a method of processing a video including baseball comprising identifying a plurality of segments of the video, where each of the segments includes a plurality of frames of the video, based upon a series of activities defined by the rules of baseball (pp.871-873, sections 1.1, 1.2, 2.1 and 2.2; series of activities such as scenes in which a batter was struck out or got a hit or a home run is defined by the rules of baseball using a spotting technique comprising a search of the minimal warp function by comparing input video sequence with pitching/batting model sequences) that could potentially result in at least one of a score, preventing a score, and creating a summarization of the video by including the plurality of segments where the summarization includes fewer frames than the video (Abstract; pg. 872, section 1.2; i.e. the indexed video segments is a digest of the game or summary of the video, a.k.a. compressed play).

Although Kawashima does not explicitly teach the frames are based upon a series of activities defined by the rules of sumo, Kawashima teaches the frames are based upon a series of activities defined by the rules of baseball. As previously mentioned, it would have been obvious to one of ordinary skill in the art to apply the video summarization of identified segments based upon baseball events taught by Kawashima to segments based upon other sporting events such as a sumo wrestling match. One would have been motivated to make such a combination in order to generate a compact representation of the content of video material, allowing easy browsing,

filtering, indexing and retrieval, etc., enabling users only interested in the exciting highlights of a game to skip the long and often boring portions of watching a sporting match in its entirety.

Referring to claim 9, Kawashima teaches wherein the activities are determined based upon the color characteristics of the video (pp. 872- 873, section 2. 1.3., activities are spotted by calculating the value from the count of pixels whose intensity change in successive frames are larger than a threshold wherein pixels are painted/colored to form an image produced on the screen).

Referring to claim 10, Kawashima teaches wherein the activities are determined based upon scene changes (pp. 872-873; section 1.1-2.1.4; wherein an activity such as an at bat activity is a period from a basic scene to the next basic scene).

Referring to claim 11, Kawashima teaches a method of processing a video including baseball comprising:

(a) identifying a plurality of segments of the video based upon detecting a play of the baseball game, wherein the identifying includes detecting the start of the play and detecting the end of the play, where each of the segments includes a plurality of frames of the video (pp. 871- 873, sections 1.1, 1.2, 2.1 and 2.2; e.g. detecting the start of the play in which a batter was struck out or got a hit or a home run is defined by the rules of baseball using a spotting technique comprising a search of the minimal warp function by comprising input video sequence with pitching/batting model sequences); and

(b) creating a summarization of the video by including the plurality of segments, where the summarization includes fewer frames than the video (Abstract; pg. 872, section 1.2, i.e. the

indexed video segments is a digest of the game or summary of the video, a.k.a. compressed play).

Although Kawashima does not explicitly teach the identified segments of the video being based upon detecting a play of a sumo match, Kawashima teaches the identified segments of the video being based upon detecting a play of a baseball game. As previously mentioned, it would have been obvious to one of ordinary skill in the art to apply the video summarization of identified segments based upon baseball events taught by Kawashima to segments based upon other sporting events such as a sumo wrestling match. One would have been motivated to make such a combination in order to generate a compact representation of the content of video material, allowing easy browsing, filtering, indexing and retrieval, etc., enabling users only interested in the exciting highlights of a game to skip the long and often boring portions of watching a sporting match in its entirety.

Referring to claim 12, Kawashima teaches a method of processing a video including baseball wherein the detecting the end of the play is based upon detecting the start of the play (pp.872-873; section 1.1-2.1.4; wherein a play such as an at bat activity is a period from an end of a basic scene to the start of the next basic scene).

Referring to claim 13, Kawashima teaches wherein the summarization identifies the plurality of segments of the video (pg. 872, section 1.2).

Referring to claim 14, Kawashima teaches wherein the summarization is a summarized video comprising the plurality of segments excluding at least a portion of the video other than the plurality of segments (pg. 872, section 1.2).

Referring to claim 36, Kawashima teaches a method of processing a video including baseball comprising:

(a) identifying a plurality of segments of the baseball video, wherein the identifying for the end of at least one of the segments is based upon detecting a scene change, where each of the segments includes a plurality of frames of the video (pp. 871-873, sections 1.1, 1.2, 2.1 and 2.2; wherein an activity such as an at bat activity is a period from a basic scene to the next basic scene); and

(b) creating a summarization of the video by including the plurality of segments, where the summarization includes fewer frames than the baseball video (Abstract: pg. 872, section 1.2; i.e. the indexed video segments is a digest of the game or summary of the video, a.k.a. compressed play).

Although Kawashima does not explicitly teach the identified segments of the video being of a sumo video, Kawashima teaches the identified segments of the video being of a baseball game video. As previously mentioned, it would have been obvious to one of ordinary skill in the art to apply the video summarization of identified segments based upon baseball events taught by Kawashima to segments based upon other sporting events such as a sumo wrestling match. One would have been motivated to make such a combination in order to generate a compact representation of the content of video material, allowing easy browsing, filtering, indexing and retrieval, etc., enabling users only interested in the exciting highlights of a game to skip the long and often boring portions of watching a sporting match in its entirety.

Referring to claim 37, Kawashima teaches the scene change is based upon a threshold between at least two frames (pp.872, sections 2.1.1; detecting scenes by determining a similarity compared to a threshold level between a set of frames).

15. Claim 8 is rejected under 35 U.S.C. 103(a) as being unpatentable over "Indexing of Baseball Telecast for Content-based Video Retrieval", Kawashima et al. (hereinafter Kawashima), as applied to claim 7 above, and "Automatically Extracting Highlights for TV Baseball Programs", Rui et al. (hereinafter Rui).

Referring to claim 8, Kawashima teaches all of the limitations as applied to claim 7 above. In addition, Kawashima teaches a method of processing a video including baseball wherein the summarization of the plurality of segments comprises a plurality of segments within the video (pg. 872, section 1.2; the indexed video segments of the summarization of the plurality of segments is stored as a digest of the game). Kawashima fails to explicitly teach the summarization of the plurality of segments to be in the same temporal order as the plurality of segments within the video. Rui further teaches wherein the summarization of the plurality of segments is in the same temporal order as the plurality of segments within the video (Abstract; section 5.4; Introduction; a method of allowing users to watch just the highlights of the exciting portions instead of the whole game due to time constraints, i.e. highlights are extracted automatically so that viewing time can be reduced). Therefore, it would have been obvious to one of ordinary skill in the art, having the teachings of Kawashima and Rui before him at the time the invention was made, to include Rui's method of processing a video including baseball wherein the summarization of the plurality of segments is in the same temporal order as the

plurality of segments within the video to Kawashima's method of processing a video including baseball wherein the summarization of the plurality of segments comprises a plurality of segments within the video. One would have been motivated to make such combination so that the time in which sequential plays in a game is being viewed is reduced.

16. Claims 18-21, 23-26, 29-30 and 34 are rejected under 35 U.S.C. 102(b) as being anticipated by "Event Detection and Summarization in Sports Video", Li et al. (hereinafter Li) and Standridge et al. (hereinafter Standridge) U.S. Publication 2002/0141619.

Referring to claim 18, Li teaches a method comprising identifying a plurality of segments of a video, wherein the start of the plurality of segments is identified based upon a frame of the video having spatial regions that differ in color, where each of the segments includes a plurality of frames of the video (Case Study 3: Japanese Sumo Wrestling; detecting a start scene based upon frames containing spatial regions with darker and lighter color, i.e. blobs of skin color on a stage color); and creating a summarization of the video by including the plurality of segments, where the summarization includes fewer frames than the video (Abstract; concatenating detected plays to generate a compact summarization of the video). Although Li teaches frames with regions that differ in color, Li fails to explicitly teach a frame of the video having an upper spatial region being substantially darker than a lower spatial region of the frame. Standridge teaches a frame of the video having an upper spatial region being substantially darker than a lower spatial region of the frame (page 5, paragraph 0052; comparing frames of a video that has a region that is darker, i.e. a dark or black area, in color than another region, i.e. a light or white area). It would have been obvious to one of ordinary skill in the art, having the teachings of Li

and Standridge before him at the time the invention was made, to include Standridge's method of identification of frame segments with regions of darker and lighter colors to Li's method of identification of video segments to produce a summarization. One would have been motivated to make such a combination in order to allow quick and easy distinguishing between regions of video frames that represent different objects.

Referring to claim 19, Li teaches wherein the lower spatial region comprises, at least in part, a pair of regions having a dominant color description representative of skin color (Case Study 3: Japanese Sumo Wrestling; two symmetrically distributed blobs of skin color).

Referring to claims 20 and 21, Li teaches the lower spatial region comprises, at least in part, a pair of regions having a dominant color description representative of stage color (Case Study 3: Japanese Sumo Wrestling; two symmetrically distributed blobs of skin color on a stage, whose color is relatively fixed).

Referring to claims 23-25, Li teaches identifying the start of plurality of segments based upon a pair of regions having a dominant color description (Case Study 3: Japanese Sumo Wrestling; detecting a start scene based upon frames containing blobs of skin color). Applicant has not disclosed that varying the percentage values of the pair of regions included in the dominate color description provides an advantage, is used for a particular purpose, or solves a stated problem. One of ordinary skill in the art, furthermore, would have expected Applicant's invention to perform equally well with the dominant color description including a percentage of the pair of regions as shown in Figure 7 because limitations of varying percentage values of the pair of regions included in the dominate color description are design choices that do not affect the functionality of the method of identifying a plurality of segments based upon a pair of

regions having a dominant color description. Therefore, at the time the invention was made, it would have been obvious to one of ordinary skill in the art, to modify the dominant color descriptions taught by Li to include varying percentages of the pair of regions, such as 25 percent, 50 percent, or 75 percent to obtain the invention as specified in claims. One would have been motivated to make such a combination in order to provide users with a plurality of implementation preferences to choice from.

Referring to claim 26, Li teaches identifying the start of plurality of segments based upon a pair of regions having a dominant color description (Case Study 3: Japanese Sumo Wrestling; detecting a start scene based upon frames containing blobs of skin color). Applicant has not disclosed that the position of the pair of regions in a particular portion of the video provides an advantage, is used for a particular purpose, or solves a stated problem. One of ordinary skill in the art, furthermore, would have expected Applicant's invention to perform equally well with the pair of regions, or blobs to be placed anywhere in the video such as shown in Figure 7, because limitations of positional placement of the pair of regions are design choices that do not affect the functionality of the method of identifying a plurality of segments based upon a pair of regions having a dominant color description. Therefore, at the time the invention was made, it would have been obvious to one of ordinary skill in the art, to modify the pair of regions, or skin color blobs taught by Li to be placed anywhere in the video, such as the lower portion to obtain the invention as specified in claims. One would have been motivated to make such a combination in order to provide users with a plurality of implementation preferences to choice from.

Referring to claims 29-30, Li teaches identifying a plurality of segments a video, wherein the start of the plurality of segments is identified based upon a pair of regions generally

symmetric to each other with respect to a generally center column of a frame of the video, where each of the segments includes a plurality of frames of the video (Case Study 3: Japanese Sumo Wrestling; detecting a start scene based upon frames containing two symmetrically distributed regions, or blobs of skin color). Applicant has not disclosed that varying the percentage values of the center column within the center of the frame provides an advantage, is used for a particular purpose, or solves a stated problem. One of ordinary skill in the art, furthermore, would have expected Applicant's invention to perform equally well with the symmetric blobs to be symmetric to each other with respect to the center column between the two blobs shown in Figure 7 and recited in Case Study 3 because limitations of the center column being within a varying percentage value within the center of the frame are design choices that do not affect the functionality of the method. Therefore, at the time the invention was made, it would have been obvious to one of ordinary skill in the art, to modify the center of the two symmetric blobs to be within any number of varying percentages of the center of the frame to obtain the invention as specified in the claims. One would have been motivated to make such a combination in order to provide users with a plurality of implementation preferences to choice from.

Referring to claim 34, Li teaches a method comprising identifying a plurality of segments of the video, wherein the start of the plurality of segments is identified based upon a frame of the video having spatial regions that differ in color (Case Study 3: Japanese Sumo Wrestling; detecting a start scene based upon frames containing spatial regions with different colors, i.e. blobs of skin color on a stage color), wherein a lower spatial regions comprises, at least in part, a pair of regions having a dominant color description representative of skin tone (Case Study 3: Japanese Sumo Wrestling; two symmetrically distributed blobs of skin color), wherein a lower

spatial region comprises, at least in part, the pair of regions having a dominant color description representative of stage color (Case Study 3: Japanese Sumo Wrestling; two symmetrically distributed blobs of skin color on a stage, whose color is relatively fixed), wherein the pair of regions are generally symmetric to each other with respect to a generally center column of a frame of the video (Case Study 3: Japanese Sumo Wrestling; detecting a start scene based upon frames containing two symmetrically distributed regions, or blobs of skin color), wherein the pair of regions move toward one another (Case Study 3: Japanese Sumo Wrestling; detecting a start scene based upon frames containing two symmetrically distributed blobs of skin color that converge), where each of the segments includes a plurality of frames of the video (Section 3. Detection of the Plays; a segment representing a play includes a sequence of shots, or frames); and creating a summarization of the video by including the plurality of segments, where the summarization includes fewer frames than the video (Abstract; concatenating detected plays to generate a compact summarization of the video). Although Li teaches frames with regions that differ in color, Li fails to explicitly teach a frame of the video having an upper spatial region being substantially darker than a lower spatial region of the frame. Standridge teaches a frame of the video having an upper spatial region being substantially darker than a lower spatial region of the frame (page 5, paragraph 0052; comparing frames of a video with a region that is darker, i.e. a dark or black area, in color than another region, i.e. a light or white area). It would have been obvious to one of ordinary skill in the art, having the teachings of Li and Standridge before him at the time the invention was made, to include Standridge's method of identification of frame segments with regions of darker and lighter colors to Li's method of identification of video segments to produce a summarization. One would have been motivated to make such a

combination in order to allow quick and easy distinguishing between regions of video frames that represent different objects.

17. Claims 35 and 42 are rejected under 35 U.S.C. 103(a) as being unpatentable over "Indexing of Baseball Telecast for Content-based Video Retrieval", Kawashima et al. (hereinafter Kawashima) and "Multimedia Content Analysis", Wang et al. (hereinafter Wang).

Referring to claim 35, Kawashima teaches a method comprising: identifying a plurality of segments of the video based upon an event, wherein the identifying for at least one of the segments includes detecting the start of the segment based upon processing of a first single frame of the video, where each of the segments includes a plurality of frames of the video (pp. 871-873, sections 1.1, 1.2, 2.1 and 2.2); and creating a summarization of the video by including the plurality of segments, where the summarization includes fewer frames than the video (Abstract; pg. 872, section 1.2; i.e. the indexed video segments are a digest of the game or summary of the video, a.k.a. compressed play). Kawashima fails to explicitly teach verifying that the first single frame is an appropriate start of the segment based upon processing of another single frame temporally relevant to the first single frame. Wang teaches verifying that the first single frame is an appropriate start of the segment based upon processing of another single frame temporally relevant to the first single frame (pg. 22, left column, lines 1-8). Therefore, it would have been obvious to one of ordinary skill in the art, having the teachings of Kawashima and Wang before him at the time the invention was made, to include Wang's verifying that the first single frame is an appropriate start of the segment based upon processing of another single frame temporally relevant to the first single frame to Kawashima's start of the segment based upon processing of

another single frame temporally relevant to the first single frame. One would have been motivated to make such a combination in order to reduce errors in segmenting related scenes.

Referring to claim 42, as best understood by the examiner, Kawashima teaches a method comprising identifying a plurality of segments of the video wherein each of the segments includes a plurality of frames of the video (Abstract) and creating a summarization of the video by including the plurality of segments, where the summarization includes fewer frames than the video (Abstract; pg. 872, section 1.2; i.e. the indexed video segments is a digest of the game or summary of the video, a.k.a. compressed play). Kawashima fails to explicitly teach detecting a segment that has a temporally sufficiently short duration , or separating/removing the identified segment from a summarization: Wang teaches a method of processing a video comprising identifying a segment that has a temporally sufficiently short duration , or separating/removing the identified segment from a summarization (pg. 21, right column; pg. 29, right column; separation of interested video portions and commercials). Therefore, it would have been obvious to one of ordinary skill in the art, having the teachings of Kawashima and Wang before him at the time the invention was made, to include Wang's method of detecting segments that has a temporally sufficiently short duration and separating/removing the identified segment from a summarization to Kawashima's method of detecting a play of the baseball game. One would have been motivated to make such a combination in order to provide users with additional criteria in content based video retrieval.

18. Claims 38 is rejected under 35 U.S.C. 103(a) as being unpatentable over "Indexing of Baseball Telecast for Content-based Video Retrieval", Kawashima et al. (hereinafter

Kawashima), as applied to claim 36 above, and “Performance Characterization of Video-Shot-Change Detection Methods”, Gargi et al. (hereinafter Gargi).

Referring to claim 38, Kawashima teaches all the limitations as applied to claim 36 above, and scene changes based upon a threshold level (pp.872, sections 2.1.1; detecting scenes by determining a similarity compared to a threshold level between a set of frames). However, Kawashima fails to explicitly teach the scene change based upon a gradual transition below a threshold level. Gargi teach detecting scene changes in a video similar to that of Gargi et al. In addition, Gargi further teach detecting shot changes based upon gradual transitions (Introduction). It would have been obvious to one of ordinary skill in the art, having the teachings of Kawashima and Gargi before him at the time the invention was made, to include Gargi’s method of detecting scene changes based upon gradual transitions to Kawashima’s method of detecting scene changes based upon a threshold level. One would have been motivated to make such a combination in order to provide users with an implementation preference.

19. Claims 39-41 and 43-44 are rejected under 35 U.S.C. 103(a) as being unpatentable over “Indexing of Baseball Telecast for Content-based Video Retrieval”, Kawashima et al. (hereinafter Kawashima), “Multimedia Content Analysis”, Wang et al. (hereinafter Wang) and “Automatically Extracting Highlights for TV Baseball Programs”, Rui et al. (hereinafter Rui).

Referring to claim 39, Kawashima teaches a method comprising identifying a plurality of segments of the video wherein each of the segments includes a plurality of frames of the video (Abstract) and creating a summarization of the video by including the plurality of segments,

where the summarization includes fewer frames than the video (Abstract; pg. 872, section 1.2; i.e. the indexed video segments is a digest of the game or summary of the video, a.k.a. compressed play). Kawashima fails to explicitly teach identifying a plurality of segments that are temporally separated by a sufficiently short duration and then connecting the identified plurality of segments. Wang teaches a method of processing a video comprising identifying a plurality of segments that are temporally separated by a sufficiently short duration and then connecting the identified plurality of segments (pg. 21, right column; pg. 29, right column; separation of interested video portions and commercials). Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made, to include Wang's method of identifying a plurality of segments that are temporally separated by a sufficiently short duration to Kawashima's method of detecting a play of the baseball game. One would have been motivated to make such a combination in order to provide users with additional criteria in content-based video retrieval. However, the modified Kawashima still does not explicitly teach connecting the identified plurality of segments. Rui teaches a method of processing a video comprising of connecting the identified segments so that the summary is in the same temporal order as the plurality of segments within the video (Abstract; section 5.4; Introduction; a method of allowing users to watch just the highlights of the exciting portions instead of the whole game due to time constraints, i.e. highlights are extracted automatically so that viewing time can be reduced). Therefore, it would have been obvious to one of ordinary skill in the art, having the teachings of modified Kawashima and Wang before him at the time the invention was made, to include Rui's method of processing a video including baseball comprising of connecting the identified plurality of segments to the modified Kawashima's method of processing a video

including baseball comprising of a plurality of segments within the video. One would have been motivated to make such a combination so that the time in which sequential plays in a game is being viewed is reduced.

Referring to claims 40-41 and 43-44, as best understood by the examiner, the modified Kawashima teaches a method, wherein the connecting includes discarding the frames of the video between the identified plurality of segments and wherein the connecting results in a single segment that includes the identified plurality of segments together with the frames of the video between the identified plurality of segments (Wang: pg. 21, right column; pg. 29, right column; separation of interested video portions and commercials; Rui: Abstract; section 5.4; Introduction; a method of allowing users to watch just the highlights of the exciting portions instead of the whole game due to time constraints, i.e. highlights are extracted automatically so that viewing time can be reduced).

20. Claim 45 is rejected under 35 U.S.C. 103(a) as being unpatentable over "Indexing of Baseball Telecast for Content-based Video Retrieval", Kawashima et al. (hereinafter Kawashima) and "Detection of Slow-Motion Replay Segments in Sports Video for Highlights Generation", Pan et al. (hereinafter Pan).

Referring to claim 45, Kawashima teaches a method comprising identifying a plurality of segments of the video wherein each of the segments includes a play of baseball wherein the segments include full-speed plays and creating a summarization of the video by including the plurality of segments, where the summarization includes fewer frames than the video, where a user may select from the summarization including only full-speed plays (Abstract; pg. 872,

section 1.2; i.e. the indexed video segments is a digest of the game or summary of the video, a.k.a. compressed play where users may select a full-speed play segment among the plurality of segments). Kawashima fails to explicitly teach segments that include slow motion plays of the full-speed plays and creating a summarization where a user may select from the summarization comprising only of slow motion plays. Pan teaches a method of process a video including baseball comprising identifying a plurality of segments of the video wherein each of the segments includes a play of baseball (“Introduction”, left column) wherein the segments include slow motion plays of the full-speed plays (“Introduction”, right column; in processing the video, slow motion plays of the full-speed plays and full-speed plays are identified) and users may select from the summarization comprising only of slow motion plays. It would have been obvious to one of ordinary skill in the art, having the teachings of Kawashima and Pan before him at the time the invention was made, to include Pan’s segments that include slow-motion plays of the full-speed plays and creating a summarization where a user may select from the summarization comprising only of slow motion plays to Kawashima’s segments that include full-speed plays and creating a summarization where a user may select from the summarization comprising only of full-speed plays. One would have been motivated to make such a combination in order to provide users with the ability to capture inherently important events. Although Kawashima and Pan do not explicitly teach the identified segments of the video including a play of sumo, Kawashima and Pan teach the identified segments of the video including a play of baseball. As previously mentioned, it would have been obvious to one of ordinary skill in the art to apply the video summarization of identified segments based upon baseball events taught by Kawashima to segments based upon other sporting events such as a

sumo wrestling match. One would have been motivated to make such a combination in order to generate a compact representation of the content of video material, allowing easy browsing, filtering, indexing and retrieval, etc., enabling users only interested in the exciting highlights of a game to skip the long and often boring portions of watching a sporting match in its entirety.

21. Claims 46-47 are rejected under 35 U.S.C. 103(a) as being unpatentable over "Automatically Extracting Highlights for TV Baseball Programs", Rui et al. (hereinafter Rui).

Referring to claim 46, Rui teaches a method of processing a video including baseball comprising identifying a plurality of segment of the video wherein each of the segments includes a play of baseball, creating a summarization of the video by including the plurality of segments wherein the summarization includes fewer frames than the video (Abstract) and removing at least one of the segments from the summary based, at least in part, upon audio information related to the at least one of the segments (pg. 105, right column, lines 24-33). Although Rui does not explicitly teach the identified segments of the video include a play of sumo, Rui teaches the identified segments of the video includes a play of baseball. As previously mentioned, it would have been obvious to one of ordinary skill in the art to apply the video summarization of identified segments based upon baseball events taught by Rui to segments based upon other sporting events such as a sumo wrestling match. One would have been motivated to make such a combination in order to generate a compact representation of the content of video material, allowing easy browsing, filtering, indexing and retrieval, etc., enabling users only interested in the exciting highlights of a game to skip the long and often boring portions of watching a sporting match in its entirety.

Referring to claim 47, Rui teaches wherein the audio information is obtained exclusively from a temporal analysis (Abstract; pg. 105, right column, lines 24-33; 3rd paragraph, pg. 107 – 3.1.5 Summary, pg. 108; using audio-track features, highlights of exciting portions of a baseball video is obtained, so that users can skip the boring parts thereby reducing the viewing time).

22. Claims 48-49 are rejected under 35 U.S.C. 103(a) as being unpatentable over “Automatically Extracting Highlights for TV Baseball Programs”, Rui et al. (hereinafter Rui) and “Multimedia Content Analysis”, Wang et al. (hereinafter Wang).

Referring to claim 48, Rui teaches a method comprising identifying a plurality of segments of the video wherein each of the segments includes a play of baseball, creating a summarization of the video by including the plurality of segments wherein the summarization includes fewer frames than the video and the duration of at least one of the segments from the summary is based, at least in part, upon audio information related to the at least one of the segments (Abstract; pg. 105, right column, lines 24-33). However, Rui fails to explicitly teach modifying the duration of at least one of the segments from the summary based, at least in part, upon audio information related to the at least one of the segments. Wang teaches a method of processing a video comprising identifying a plurality of segments of the video, creating a summarization of the video by including the plurality of segments wherein the summarization includes fewer frames than the video and modifying the duration of at least one of the segments from the summary based, at least in part, upon audio information related to the at least one of the segments (pg. 29, left column, lines 49-53; pg. 30, left column, lines 6-22). It would have been obvious to one of ordinary skill in the art, having the teachings of Rui and Wang before him at

the time the invention was made, to include Wang's modifying the duration of at least one of the segments from the summary based, at least in part, upon audio information related to the at least one of the segments to the method of Rui wherein the duration of at least one of the segments from the summary based, at least in part, upon audio information related to the at least one of the segments. One would have been motivated to make such a combination in order to provide users with a more customized method of processing a video. Although Rui does not explicitly teach the identified segments of the video include a play of sumo, Rui teaches the identified segments of the video includes a play of baseball. As previously mentioned, it would have been obvious to one of ordinary skill in the art to apply the video summarization of identified segments based upon baseball events taught by Kawashima to segments based upon other sporting events such as a sumo wrestling match. One would have been motivated to make such a combination in order to generate a compact representation of the content of video material, allowing easy browsing, filtering, indexing and retrieval, etc., enabling users only interested in the exciting highlights of a game to skip the long and often boring portions of watching a sporting match in its entirety.

Referring to claim 49, the modified Rui teaches wherein the audio information is obtained exclusively from a temporal analysis (Rui: Abstract; pg. 105, right column, lines 24-33; 3rd paragraph, pg. 107 - 3.1.5 Summary, pg. 108; using audio-track features, highlights of exciting portions of a baseball video is obtained, so that users can skip the boring parts thereby reducing the viewing time).

23. Claim 61 is rejected under 35 U.S.C. 102(b) as being anticipated by “Event Detection and Summarization in Sports Video”, Li et al. (hereinafter Li) and “Indexing of Baseball Telecast for Content-based Video Retrieval”, Kawashima et al. (hereinafter Kawashima).

Referring to claim 61, as best understood by the examiner, Li teaches a method comprising identifying a plurality of segments of a video based upon: a pair of substantially white regions generally symmetric with respect to the center of the image (Case Study 3: Japanese Sumo Wrestling and Figure 7; a pair of blobs of skin color, i.e. a pair substantially white regions, symmetrically distributed); the image free from other significant substantially white areas (Case Study 3: Japanese Sumo Wrestling and Figure 7; the blobs of skin color, or substantially white regions are on a stage of a uniform color, i.e. the black regions shown in Figure 7); the white regions persist for a plurality of seconds (Case Study 3: Japanese Sumo Wrestling and Figure 7; the blobs persist for a plurality of seconds, i.e. converge towards one another); the white regions preceding the start of a play (Case Study 3: Japanese Sumo Wrestling and Figure 7; the blobs, or white regions, are tracked to see if they converge, to detect the real start of a play); where each of the segments includes a plurality of frames of the video (Section 3. Detection of the Plays; a segment representing a play includes a sequence of shots, or frames); and creating a summarization of the video by including the plurality of segments, where the summarization includes fewer frames than the video (Abstract; concatenating detected plays to generate a compact summarization of the video). However, Li fails to explicitly teach the detection of graphical text segments of the video. Kawashima teaches the detection of graphical text segments of the video (pp. 871, section 1.1; pp. 872, section 2.1.2; detecting texts superimposed on the scenes). It would have been obvious to one of ordinary skill in the art,

having the teachings of Li and Kawashima before him at the time the invention was made, to include Kawashima's method of detecting text segments of a video to Li's method of identifying a plurality of segments of a video. One would have been motivated to make such a combination in order to quickly and easily identify specific scenes in video, such as the beginning and end of a play, via descriptions stored and indexed with scenes or segments.

24. The prior art made of record on form PTO-892 and not relied upon is considered pertinent to applicant's disclosure. Applicant is required under 37 C.F.R. § 1.111(c) to consider these references fully when responding to this action. The documents cited therein teach similar methods of creating a summarization from identified video segments.

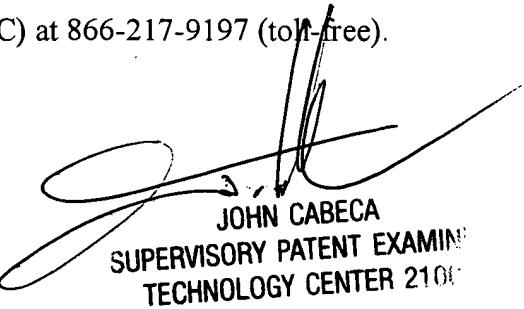
Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Ting Zhou whose telephone number is (571) 272-4058. The examiner can normally be reached on Monday - Friday 7:00 am - 4:30 pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, John Cabeca can be reached at (571) 272-4048. The fax phone number for the organization where this application or proceeding is assigned is (571) 273-4058.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

TZ



JOHN CABEZA
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2100